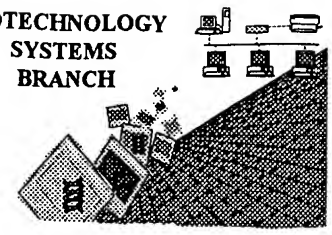
0590
1115

BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/002,631A
Source: OIPE
Date Processed by STIC: 11/21/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm>, **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
   Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, r other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/002,631A

DATE: 11/21/2002
TIME: 10:08:46

Input Set : A:\10002631SEQLISTING.txt
Output Set: N:\CRF4\11212002\J002631A.raw

pp. 1-2, 5

Does Not Comply
Corrected Diskette Needed

```
     4 <110> APPLICANT: Graff, Jonathon M.
     5       Muenster, Matthew
     7 <120> TITLE OF INVENTION: METHODS TO IDENTIFY SIGNAL SEQUENCES
    10 <130> FILE REFERENCE: A34943 (090495.0243)
    12 <140> CURRENT APPLICATION NUMBER: 10/002,631A
C--> 13 <141> CURRENT FILING DATE: 2002-10-31
    15 <150> PRIOR APPLICATION NUMBER: 60/300,309
    16 <151> PRIOR FILING DATE: 2001-06-21
    18 <160> NUMBER OF SEQ ID NOS: 324
    20 <170> SOFTWARE: FastSEQ for Windows Version 4.0
    22 <210> SEQ ID NO: 1
    23 <211> LENGTH: 884
    24 <212> TYPE: DNA
    25 <213> ORGANISM: Homo sapiens
    27 <220> FEATURE:
    28 <221> NAME/KEY: unsure
    29 <222> LOCATION: (608)...(884)
    30 <223> OTHER INFORMATION: n = A, C, G or T
    32 <400> SEQUENCE: 1
    33 ggatccagtg gcaaaaaaac aaacaacaaa caacaaacaa aacaaaacaa acaaacaaaa 60
    34 aatcccacca atcttcatgg gtaaactttc ctgctcaggg atgtaagctg actctagacc 120
    35 atctcgcggt tcctgcggat agcacagcac aagatcatac tgaagatcat gccaaatatc 180
    36 atgaccacgg caatgccgat gcccactgcg ccgatgatgt ggaatttatt gtcgaagacc 240
    37 tctttgatgg catcaggaca ggacttcacg gtgaaggttt cgagtacgtc cttcttgggg 300
    38 cagatgtctg agataaactg ttccacgccc ccagccaaac cacagcagtt caacgcatag 360
    39 tggatggctt tcagcgtttc ccgctggggc tcatccttgg ttttcagctt gttgtaggtg 420
    40 tccttgtaaa actcctggac ttccttaatc acctcatcct tgtgggaata tccccagatg 480
    41 gccgcagcta tttcaatggc gaatatcacc aagaggaagc ccgaagaaca gtcccagcat 540
    42 gcactgggac tcctgcacag ccccgcagca gcccaggaag cccaccagca tcatgagggc 600
W--> 43 gccggctncg atcagaatat agactcctgt gtagaagctg gaattattat tattaagttt 660
W--> 44 cttgctcgaa gatgctcttg gnctgagagt cgaatcggaa cccttagtca atggcaagga 720
W--> 45 cagnaattcc cgggnaaggc ccnaannaag aannttaaat cccgaacaag natggtattt 780
W--> 46 gntncccttt ggggcctncn tttntaccgg nnttttgtna nggnntnact taanccnggg 840
W--> 47 cccnaacggg ttccggnant tgggggncnc cccccnantn ngnn                  884
    49 <210> SEQ ID NO: 2
    50 <211> LENGTH: 288
    51 <212> TYPE: PRT
    52 <213> ORGANISM: Homo sapiens
    54 <220> FEATURE:
    55 <221> NAME/KEY: UNSURE
    56 <222> LOCATION: (1)...(2)
    57 <223> OTHER INFORMATION: Xaa = Any amino acid
    59 <400> SEQUENCE: 2
```

Xaa's are also shown at locations 3 through 92

(see p. 2)

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/002,631A

DATE: 11/21/2002
TIME: 10:08:46

Input Set : A:\10002631SEQLISTING.txt
Output Set: N:\CRF4\11212002\J002631A.raw

W--> 60 Xaa Xaa Xaa Gly Xaa Xaa Pro Xaa Xaa Arg Asn Pro Xaa Gly Pro Xaa
61 1 5 10 15
W--> 62 Xaa Lys Xaa Xaa Xaa Xaa Lys Xaa Pro Val Xaa Xaa Xaa Ala Pro Lys
63 20 25 30
W--> 64 Gly Xaa Lys Tyr His Xaa Cys Ser Gly Phe Xaa Xaa Leu Xaa Xaa Gly
65 35 40 45
W--> 66 Leu Xaa Arg Glu Xaa Leu Ser Leu Pro Leu Thr Lys Gly Ser Asp Ser
67 50 55 60
W--> 68 Thr Leu Xaa Pro Arg Ala Ser Ser Ser Lys Lys Leu Asn Asn Asn Asn
69 65 70 75 80
W--> 70 Ser Ser Phe Tyr Thr Gly Val Tyr Ile Leu Ile Xaa Ala Gly Ala Leu
71 85 90 95
72 Met Met Leu Val Gly Phe Leu Gly Cys Cys Gly Ala Val Gln Glu Ser
73 100 105 110
74 Gln Cys Met Leu Gly Leu Phe Phe Gly Leu Pro Leu Gly Asp Ile Arg
75 115 120 125
76 His Asn Ser Cys Gly His Leu Gly Ile Phe Pro Gln Gly Gly Asp Gly
77 130 135 140
78 Ser Pro Gly Val Leu Gln Gly His Leu Gln Gln Ala Glu Asn Gln Gly
79 145 150 155 160
80 Ala Pro Ala Gly Asn Ala Glu Ser His Pro Leu Cys Val Glu Leu Leu
81 165 170 175
82 Trp Phe Gly Trp Gly Arg Gly Thr Val Tyr Leu Arg His Leu Pro Gln
83 180 185 190
84 Glu Gly Arg Thr Arg Asn Leu His Arg Glu Val Leu Ser Cys His Gln
85 195 200 205
86 Arg Gly Leu Arg Gln Ile Pro His His Arg Arg Ser Gly His Arg His
87 210 215 220
88 Cys Arg Gly His Asp Ile Trp His Asp Leu Gln Tyr Asp Leu Val Leu
89 225 230 235 240
90 Cys Tyr Pro Gln Glu Pro Arg Asp Gly Leu Glu Ser Ala Tyr Ile Pro
91 245 250 255
92 Glu Gln Glu Ser Leu Pro Met Lys Ile Gly Gly Ile Phe Cys Leu Phe
93 260 265 270
94 Val Leu Phe Cys Leu Leu Phe Val Val Cys Phe Phe Ala Thr Gly Ser
95 275 280 285
98 <210> SEQ ID NO: 3
99 <211> LENGTH: 529
100 <212> TYPE: DNA
101 <213> ORGANISM: Homo sapiens
103 <400> SEQUENCE: 3
104 actgatcttc agcatctttt actttcacca gcgtttctgg gtgaaagaaa acattcccca 60
105 gggaagacaa aagcaacaag ctcagggctg acatcaagat acctcccaga aagaggtagc 120
106 tacggcgcct ggcatagagt gcactgaggg tgaagcaggt aaagatcatt gccgtgccca 180
107 tgaaagcagt gggaaggatg ctggggttga cagcaataca aaactccagg gcagggccca 240
108 ggccaactcc tgtaaggaat gcaaatccag caagaagtcc cagtcttttc tgttcagttt 300
109 catggctatg aggtgttgcc atcagccaaa tcatcaatat cagggagccc aaggcagaca 360
110 gcaggccagc ctgaatgaaa tgagtgacca tatggacata ggcccctgca gccgccacaa 420
111 acatacaaag ggcaaaactt gcatagacct tcttcaggtg ctgctgcgtt gacggggtta 480

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/002,631A

DATE: 11/21/2002
TIME: 10:08:46

Input Set : A:\10002631SEQLISTING.txt
Output Set: N:\CRF4\11212002\J002631A.raw

112 tatgagaaaa ttttaaaagc gcatcaaagg tcgacgcggc cgcgaattc 529
114 <210> SEQ ID NO: 4
115 <211> LENGTH: 162
116 <212> TYPE: PRT
117 <213> ORGANISM: Homo sapiens
119 <400> SEQUENCE: 4
120 Glu Phe Ala Ala Ala Ser Thr Phe Asp Ala Leu Leu Lys Phe Ser His
121 1 5 10 15
122 Ile Thr Pro Ser Thr Gln Gln His Leu Lys Lys Val Tyr Ala Ser Phe
123 20 25 30
124 Ala Leu Cys Met Phe Val Ala Ala Ala Gly Ala Tyr Val His Met Val
125 35 40 45
126 Thr His Phe Ile Gln Ala Gly Leu Leu Ser Ala Leu Gly Ser Leu Ile
127 50 55 60
128 Leu Met Ile Trp Leu Met Ala Thr Pro His Ser His Glu Thr Glu Gln
129 65 70 75 80
130 Lys Arg Leu Gly Leu Leu Ala Gly Phe Ala Phe Leu Thr Gly Val Gly
131 85 90 95
132 Leu Gly Pro Ala Leu Glu Phe Cys Ile Ala Val Asn Pro Ser Ile Leu
133 100 105 110
134 Pro Thr Ala Phe Met Gly Thr Ala Met Ile Phe Thr Cys Phe Thr Leu
135 115 120 125
136 Ser Ala Leu Tyr Ala Arg Arg Arg Ser Tyr Leu Phe Leu Gly Gly Ile
137 130 135 140
138 Leu Met Ser Ala Leu Ser Leu Leu Leu Leu Ser Ser Leu Gly Asn Val
139 145 150 155 160
140 Phe Phe
144 <210> SEQ ID NO: 5
145 <211> LENGTH: 454
146 <212> TYPE: DNA
147 <213> ORGANISM: Homo sapiens
149 <400> SEQUENCE: 5
150 ggatccgggc caaaaaaaat aaacagcaac ttcatagaca aaaaaggaaa aaaaaagaaa 60
151 ccttttatct ttggcctttt taaccatctc atacaaacca actacttata gtacagctaa 120
152 gtacatacac aaaaaagtta ctggaatgct cggaataaga ttgtttttct gttgtcattt 180
153 ttgctttttt tacaaggttt ttttttctcct ttgagattat aatgaacatg gtcacaccac 240
154 aagtaaagtc agaagtagga cagagaacgc tccgaaggct ggtttggtca tccgagatca 300
155 ttaaaaatgg ctgaccctaa caatatgtac aaaaatataa aatgtaaata aaaaatacaa 360
156 acaaatttcc tttttaaagt actttaagaa aaaaagcagg gccttggaag ttttggttct 420
157 tttttcctcc cctggtcgac gcggccgcga attc 454
159 <210> SEQ ID NO: 6
160 <211> LENGTH: 144
161 <212> TYPE: PRT
162 <213> ORGANISM: Homo sapiens
164 <400> SEQUENCE: 6
165 Asn Ser Arg Pro Arg Arg Pro Gly Glu Glu Lys Arg Thr Lys Thr Ser
166 1 5 10 15
167 Lys Ala Leu Leu Phe Phe Leu Lys Tyr Phe Lys Lys Glu Ile Cys Leu
168 20 25 30

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/002,631A

DATE: 11/21/2002
TIME: 10:08:46

Input Set : A:\10002631SEQLISTING.txt
Output Set: N:\CRF4\11212002\J002631A.raw

```
169 Tyr Phe Leu Phe Thr Phe Tyr Ile Phe Val His Ile Val Arg Val Ser
170         35                  40                  45
171 His Phe Ser Arg Met Thr Lys Pro Ala Phe Gly Ala Phe Ser Val Leu
172     50                  55                  60
173 Leu Leu Thr Leu Leu Val Val Pro Cys Ser Leu Ser Gln Arg Arg Lys
174 65                  70                  75                  80
175 Lys Thr Leu Lys Lys Gln Lys Gln Gln Lys Asn Asn Leu Ile Pro Ser
176                 85                  90                  95
177 Ile Pro Val Thr Phe Leu Cys Met Tyr Leu Ala Val Leu Val Val Gly
178             100                 105                 110
179 Leu Tyr Glu Met Val Lys Lys Ala Lys Asp Lys Arg Phe Leu Phe Phe
180         115                 120                 125
181 Ser Phe Phe Val Tyr Glu Val Ala Val Tyr Phe Phe Trp Pro Gly Ser
182     130                 135                 140
185 <210> SEQ ID NO: 7
186 <211> LENGTH: 478
187 <212> TYPE: DNA
188 <213> ORGANISM: Homo sapiens
190 <400> SEQUENCE: 7
191 ggatccaagc atcaggagca ggcaaggaga accaaaagac atcaagaaac cgatttgctt 60
192 gagaaaagca gcgattcttc ctttcagagt tctccatggc tcagaaaatg cccaagacat 120
193 catgtatgtg acttagatac tgctttttgg gaggttaaga gtagcatgaa gaacttaaga 180
194 tgacgataag agtctaaatt tttagtttca aggtttcaat agaatgtgga tatattcaaa 240
195 actttcaaaa aggacagtgt ttagaaaggg taaaactagg acacagaaaa cactgggaat 300
196 taccacgacc cccaagtgct tccggctcca ggaaataacc attcatgtgt ttgctggagg 360
197 tcacacaatt ttcccctatt acctggtgca aaatgactca tcacttccca aaagcttctt 420
198 ttcaaaccac gattttccca tttattttgg tccaatgcgt cgacgcggcc gcgaattc   478
200 <210> SEQ ID NO: 8
201 <211> LENGTH: 150
202 <212> TYPE: PRT
203 <213> ORGANISM: Homo sapiens
205 <400> SEQUENCE: 8
206 Asn Ser Arg Pro Arg Arg Arg Ile Gly Pro Lys Met Gly Lys Ser Trp
207  1               5                  10                  15
208 Phe Glu Lys Lys Leu Leu Gly Ser Asp Glu Ser Phe Cys Thr Arg Gly
209             20                  25                  30
210 Lys Ile Val Pro Pro Ala Asn Thr Met Val Ile Ser Trp Ser Arg Lys
211         35                  40                  45
212 His Leu Gly Val Val Val Ile Pro Ser Val Phe Cys Val Leu Val Leu
213     50                  55                  60
214 Pro Phe Leu Asn Thr Val Leu Phe Glu Ser Phe Glu Tyr Ile His Ile
215 65                  70                  75                  80
216 Leu Leu Lys Pro Asn Lys Phe Arg Leu Leu Ser Ser Ser Val Leu His
217                 85                  90                  95
218 Ala Thr Leu Asn Leu Pro Lys Ser Ser Ile Val Thr Tyr Met Met Ser
219             100                 105                 110
220 Trp Ala Phe Ser Glu Pro Trp Arg Thr Leu Lys Gly Arg Ile Ala Ala
221         115                 120                 125
222 Phe Leu Lys Gln Ile Gly Phe Leu Met Ser Phe Gly Ser Pro Cys Leu
```

RAW SEQUENCE LISTING DATE: 11/21/2002
PATENT APPLICATION: **US/10/002,631A** TIME: 10:08:46

Input Set : **A:\10002631SEQLISTING.txt**
Output Set: **N:\CRF4\11212002\J002631A.raw**

```
223     130                 135                 140
224 Leu Leu Met Leu Gly Ser
225 145                 150
228 <210> SEQ ID NO: 9
229 <211> LENGTH: 770
230 <212> TYPE: DNA
231 <213> ORGANISM: Homo sapiens
233 <220> FEATURE:
234 <221> NAME/KEY: unsure
235 <222> LOCATION: (615)...(757)
236 <223> OTHER INFORMATION: n = A, C, G or T
238 <400> SEQUENCE: 9
239 ggatcctgct gtgttggtct ggtagcttcg gctgctgtaa gtgacaagtt gtagttgcct 60
240 gttgagttgg tccagccctg ggctgacaag ggtgagatct gcctgaccct ctccagtgag 120
241 agtaactcca gtcacttccc ctgccacgtc ccaggtgcct agggaggcag tcaggttcac 180
242 ctggtatacc tcctgaccag aagctgcctg aaggctcagc cctggcacca agatgctcct 240
243 gaggggctga acttccacac cctgtagggg gtactggagc ggggagttgg caggggctat 300
244 gagcagctgg tcagctgggg actggctcct cgacagaaag gcctggaact cctgctctct 360
245 tgtggcagag gcagccctca gctctgcagg gtcaaaggcc ttggtgaggt caatagctcg 420
246 gacttgtttc tggaagggga gggggaggcc ccccccactg gactcacaac tgcagttgtt 480
247 ccaaggccagc agccccacta cttgctcctt gatcctgacc gggatgtgtg cctagcgggg 540
W--> 248 ctcangagca agatctggca gctcgggcct gcgggggctt tgcgggggcg cccacggcgc 600
W--> 249 aagaagtacc cggangcccg ggcgccgtnc cgggtgctcg cgtacaggan ccccancgag 660
W--> 250 gccaagccna ccagaaggac caaaacgcac aagggcccgg cgggccaacc acatcctgct 720
W--> 251 aacctntaag gacggcaaaa ttcggnccgg ctntnanccg gccggaatta          770
253 <210> SEQ ID NO: 10
254 <211> LENGTH: 255
255 <212> TYPE: PRT
256 <213> ORGANISM: Homo sapiens
258 <220> FEATURE:
259 <221> NAME/KEY: UNSURE
260 <222> LOCATION: (5)...(75)
261 <223> OTHER INFORMATION: Xaa = Any amino acid
263 <400> SEQUENCE: 10
W--> 264 Ile Pro Ala Gly Xaa Xaa Pro Xaa Arg Ile Leu Pro Ser Leu Xaa Val
265  1               5                  10                  15
266 Ser Arg Met Trp Leu Ala Arg Arg Ala Leu Val Arg Phe Gly Pro Ser
267             20                  25                  30
W--> 268 Gly Xaa Leu Gly Leu Xaa Gly Xaa Pro Val Arg Glu His Pro Xaa Arg
269         35                  40                  45
W--> 270 Arg Pro Gly Xaa Arg Val Leu Leu Ala Pro Trp Ala Pro Pro Gln Ser
271     50                  55                  60
W--> 272 Pro Arg Arg Pro Glu Leu Pro Asp Leu Ala Xaa Glu Pro Arg Ala His
273 65                  70                  75                  80
274 Ile Pro Val Arg Ile Lys Glu Gln Val Val Gly Leu Leu Ala Trp Asn
275                 85                  90                  95
276 Asn Cys Ser Cys Glu Ser Ser Gly Gly Gly Leu Pro Leu Pro Phe Gln
277             100                 105                 110
278 Lys Gln Val Arg Ala Ile Asp Leu Thr Lys Ala Phe Asp Pro Ala Glu
```

what about "n" at location 545?
(see below)

FYI **Use of n and/or Xaa has been detected in the Sequence Listing.** Review the Sequence Listing to insure a corresponding explanation is presented in the <220> to <223> fields of each sequence using n or Xaa.